

Алгоритмизация процессов обнаружения уязвимостей интерфейсов БТС на основе вероятностных автоматов

А.В. Скатков, А.А. Брюховецкий, Д.В. Моисеев

Севастопольский государственный университет

Аннотация. Рассматривается алгоритмический подход, базирующийся на методах адаптивной интеллектуальной технологии контроля состояния объектов вычислительных систем. Подход ориентирован на обнаружение изменения состояния контролируемых ресурсов БТС: канал связи, процессор, память. При этом скорость и достоверность оценки ситуации может иметь решающее значение. Реализация таких задач в реальном времени не всегда возможна с помощью аналитического подхода, поскольку эти задачи характеризуются противоречивостью, нелинейностью, недифференцируемостью, многоэкстремальностью, сложной топологией области допустимых значений, высокой вычислительной сложностью оптимизируемых функций, высокой размерностью пространства поиска и т.п. В условиях дефицита априорной информации большая часть проблем анализа данных связана с исследованиями стохастических систем. Одним из наиболее эффективных инструментов моделирования сложных стохастических систем является методология вероятностно-автоматного моделирования. Представлена адаптивная модель с использованием байесовского классификатора оценивания изменения состояний ресурсов БТС. Модель базируется на основе вероятностного автомата с адаптивной самонастройкой.

Ключевые слова: вероятностный автомат, байесовский классификатор, динамическое оценивание ресурсов, адаптивная модель, самонастройка

Algorithmization of processes for detecting vulnerabilities of BTS interfaces based on probabilistic automata

A.V. Skatkov, A.A. Bryukhovetsky, D.V. Moiseev

Sevastopol State University

Abstract. The algorithmic approach based on the methods of adaptive intelligent technology for monitoring the state of objects of computer systems is considered. The approach is focused on the detection of changes in the state of controlled BTS

resources: communication channel, processor, memory. Moreover, the speed and reliability of the assessment of the situation can be crucial. Realization of such tasks in real time is not always possible using an analytical approach, since these tasks are characterized by inconsistency, nonlinearity, non-differentiability, multi-extremity, complex topology of the range of admissible values, high computational complexity of optimized functions, high dimension of the search space, etc. In the context of a lack of a priori information, most of the problems of data analysis are associated with studies of stochastic systems. One of the most effective tools for modeling complex stochastic systems is the methodology of probabilistic automaton modeling. An adaptive model using the Bayesian classifier for assessing changes in the state of BTS resources is presented. The model is based on a probabilistic automaton with adaptive self-tuning.

Keywords: probabilistic automaton, Bayesian classifier, dynamic resource estimation, adaptive model, self-tuning.

1. Введение

Предлагаемый в статье подход ориентирован на решение задач обнаружения моментов времени изменения состояния контролируемых объектов БТС, каковыми являются ресурсы: канал связи, процессор, память. При этом скорость и достоверность оценки ситуации может иметь решающее значение. Реализация таких задач в реальном времени не всегда возможна с помощью аналитического подхода, поскольку эти задачи характеризуются противоречивостью, нелинейностью, недифференцируемостью, многоэкстремальностью, сложной топологией области допустимых значений, высокой вычислительной сложностью оптимизируемых функций, высокой размерностью пространства поиска и т.п. В условиях дефицита априорной информации большая часть проблем анализа данных связана с исследованиями стохастических систем [1-4]. Одним из наиболее эффективных инструментов моделирования сложных стохастических систем является методология вероятностно-автоматного моделирования [5]. Продуктивность названной методологии обуславливают следующие ее особенности:

- наличие средств, обеспечивающих адекватное описание сложных стохастических систем и процессов их функционирования;
- возможность построения унифицированных моделей для широкого класса систем;
- использование систем поддержки принятия решений при необходимости в точной оценке выбора различных альтернатив на основе вероятностных автоматов.

Настоящая работа посвящена применению модели вероятностных автоматов для поддержки принятия решений при оценке информационных состояний объектов транспортной инфраструктуры умного города. К числу таких объектов, например, относятся интеллектуальные системы управления беспилотными воздушными и наземными транспортными средствами, системы

обеспечивающие межмашинное взаимодействие с использованием технологии интернет вещей и др. Разнородность приложений и беспроводных коммуникаций в инфраструктуре умного города существенно усложняет обеспечение безопасности объектов [6]. Методы предупреждения атак для безопасной эксплуатации транспортных средств должны быть динамичными и реагировать на возможные угрозы. Упреждающий подход к угрозам должен быть ключевым требованием, которое должно быть выполнено. Однако, поскольку невозможно предсказать все возможные угрозы для БТС, важно, чтобы в результате атаки у пользователей было как можно меньше нарушений. В работах [7-11] рассматриваются методы обнаружения и классификация основных типов атак на БТС, которые влияют на работоспособность транспортных средств. К числу основных типов атак относятся: распределенные атаки типа «отказ в обслуживании» (*DDoS*), атаки типа «черной дыры» (*Black-hole*), атаки «посредника» (*Man in the middle (MITM)*), *Sybil* – псевдоспуфинговые атаки, атаки на основе имперсонализации (*Impersonation*), фальсифицированно-информационные атаки (*Falsified-Information*) и другие.

В настоящее время известны различные вероятностные автоматные модели. Причина разнообразия автоматных моделей объясняется широтой области их применения. Вероятностные автоматы используются в таких областях, например, как: логическое управление, математическая лингвистика, теория формальных языков, моделирование поведения человека, при описании моделей информационной защиты предприятия и др. Критерий применимости автоматного подхода лучше всего выражается через понятие «сложное поведение» [12]. Можно сказать, что объект обладает сложным поведением, если в качестве реакции на некоторое входное воздействие он может осуществить одну из нескольких выходных реакций. При этом существенно, что реакция может зависеть не только от входного воздействия, но и от предыстории.

2. Постановка задачи.

Проводимые исследования в области интеллектуальных транспортных систем опираются на теоретический и методологический базис в областях самоорганизации сложных естественных и искусственных иммунных систем, которые обеспечивают сбалансированную стратегию нахождения решения и сочетают в себе локальный и глобальный поиск решения. Авторами проекта также получены некоторые результаты, в работах, посвященных решению задач обнаружения вторжений в телекоммуникационных сетях на основе ИИС [2-4].

Основные методы обнаружения вторжений представлены на рис. 1 [3]:

Установлено, что решение задачи обнаружения уязвимостей БТС характеризуется многомерностью, многокритериальностью, влиянием форм представления информации на точность классификации, необходимостью использования минимальной априорной информации, сочетанием детерминизма и нечеткости, возможностью сочетания формальных методов и учета экспертных суждений [10-11]. В настоящее время большая часть проблем анализа данных связана с исследованиями стохастических динамических систем, в которых обнаружение существенных, но редких информационных ситуаций часто имеет решающее значение [5].



Рис. 1. Основные методы обнаружения вторжений

Это направление соответствует развитию классического подхода к распознаванию ситуаций и построения систем поддержки принятия решений с использованием вероятностных моделей, поскольку внешняя среда является стохастической. Выявлена необходимость построения информационной технологии, так как аналитическое решение в заданных условиях невозможно. Целью работы является разработка адаптивной модели с использованием байесовского классификатора оценивания состояний ресурсов БТС. Модель базируется на основе вероятностного автомата с адаптивной самонастройкой. На основе предлагаемой модели решается задача оценки

состояний ресурсов с целью повышения достоверности результатов классификации информационных ситуаций. Обозначим

$$R = \{R_1, \dots, R_j, \dots, R_r\}$$

множество контролируемых ресурсов БТС. Определим контролируемые характеристики ресурсов, по которым будем оценивать их состояние (значения характеристик нормированы, определены в диапазоне $[0;1]$):

D_j – нагрузка j -го ресурса,

V_j – скорость изменения D_j , где $V = (D(t_i) - D(t_{i-1}))/\Delta t$.

Указанные характеристики – векторные величины с компонентами элементов множеств R_j ,

Обозначим S^t_i – состояние i -го ресурса в момент времени t . Основной задачей является адаптивная оценка значений вероятности $P(S_j)$ состояний ресурсов БТС, сформированных для нормального поведения ресурса на временном интервале T_N и полученных значений в результате внешнего воздействия на интервале T_{2N} . Применение указанной модели позволит получить достоверные оценки гипотез появления состояний S^t_i .

3 Метод обнаружения изменения состояния ресурсов БТС.

Для динамической оценки состояний ресурсов БТС предлагается использовать автоматную вероятностную модель. С этой целью определим вероятностный автомат как систему

$$\Sigma = (S, X, Y, \Phi, \Psi, S_n), (1)$$

где

$S_n \in S$ – начальное состояние,

S – множество значений вектора состояния,

X – множество значений вектора входа,

Y – множество значений вектора выхода,

Φ – функция переходов,

Ψ – функция выходов.

Будем фиксировать изменения состояния S^t_j ресурса на временном интервале T_N в моменты времени $\{t_0, t_1, \dots, t_j, \dots, t_n\}$. В соответствии с задаваемой схемой (1) автомат функционирует в дискретные моменты времени, которыми являются такты t_0, t_1, t_2, \dots . Каждому такту сопоставляется входной сигнал о состоянии ресурса – X , выходной сигнал – Y и сигнал о внутреннем состоянии – S . Будем рассматривать вероятностную автоматную модель Мура. В этом случае элементы матриц переходов и выходов представляют собой соответствующие оценки вероятностей переходов между состояниями. Введем правила для указанных переходов:

$P(s_j(t+1)) = \Phi(s_j(t), x(t))$ – вероятность перехода в новое состояние,

$P(y_j(t+1)) = \Psi(s_j(t+1))$ – вероятность появления выходного сигнала,

где $j=1, m$, m – число состояний автомата, $t = 0, 1, 2, \dots$.

В начальный момент времени t_n автомат находится в состоянии S_n . В этот момент времени выходной сигнал не вырабатывается. Для начального состояния задается распределение вероятностей перехода во внутренние состояния. Начальное распределение состояний приведено в таблице 1.

Таблица 1. Начальное распределение состояний вероятностного автомата

$\Psi(s_j(t+1))$	$Y_0(t+1)$	$Y_1(t+1)$	$Y_j(t+1)$	$Y_m(t+1)$
$\Phi(s_j(t), x(t))$	$S_0(t+1)$	$S_1(t+1)$	$S_j(t+1)$	$S_m(t+1)$
$S_n(t)$	P_{n0}	P_{n1}	P_{nj}	P_{nm}

В первом такте при поступлении сигнал x_j автомат с вероятностью P_{nj} переходит в состояние S_j и вырабатывается выходной сигнал Y_j . Начальное распределение вероятностей состояний формируется на основе априорной информации, которая может быть получена, например, в процессе нормального функционирования БТС при отсутствии внешних возмущений.

В таблице переходов задаются оценки вероятностей перехода в состояние $S_k(t+1)$ в зависимости от состояния $s_j(t)$ при условии поступления сигнала $x_i(t)$. Обозначим эту вероятность – P_{ijk} . В каждой строке матрицы вероятности перехода образуют полную группу: $\sum_{k=1}^m P(S_{i,j,k}) = 1$. Для вероятностного автомата необходимо столько таблиц переходов, сколько входных сигналов $x_i(t)$. Ниже приведен пример таблицы переходов (Таблица 2) при входном сигнале $x_i(t)$.

Таблица 2. Таблица переходов вероятностного автомата

		$S_0(t+1)$	$S_1(t+1)$	$S_j(t+1)$	$S_m(t+1)$
$x_i(t)$	$S_0(t)$	P_{i00}	P_{i01}	P_{i0j}	P_{i0m}
$x_i(t)$	$S_1(t)$	P_{i10}	P_{i11}	P_{i1j}	P_{i1m}
.....
$x_i(t)$	$S_j(t)$	P_{ij0}	P_{ij1}	P_{ijj}	P_{ijm}
.....
$x_i(t)$	$S_m(t)$	P_{im0}	P_{im1}	P_{imj}	P_{imm}

Таблица выходов автомата Мура (Таблица 3) упрощается по сравнению с автоматом Мили, так как выходной сигнал Y зависит только от внутреннего состояния S и не зависит от входного сигнала X . Обозначим P_{ij} – вероятность появления выходного состояния $Y_j(t+1)$ при условии, что автомат находился в состоянии $S_i(t+1)$.

Таблица 3. Таблица выходов вероятностного автомата Мура

$S \backslash Y$	$Y_0(t+1)$	$Y_1(t+1)$	$Y_j(t+1)$	$Y_m(t+1)$
$S_0(t+1)$	P_{00}	P_{01}	P_{0j}	P_{0m}
$S_1(t+1)$	P_{10}	P_{11}	P_{1j}	P_{1m}
.....
$S_j(t+1)$	P_{j0}	P_{jj}	P_{jj}	P_{jm}
.....
$S_m(t+1)$	P_{m0}	P_{m1}	P_{mj}	P_{mm}

При достаточно общей постановке задачи речь идет о контроле результатов наблюдений над состоянием ресурсов БТС. Будем полагать, что состояние ресурса R_j в заданный момент времени t зависит от значений характеристик D_j, V_j . Пусть эти состояния в момент времени t обозначены S^t_j – множество возможных состояний объекта R_j . Значения состояний определены и нормированы в диапазоне $[0;1]$. Пусть состояния определены на интервалах $[I_k; I_{k+1}]$, где I_k – порог для задания области определения состояния ресурса, $k=0,1,\dots,j$. На рисунке 2 приведен пример различия состояний S^t_j ресурса R_j на интервалах: $S^t_0 \in [0; I_0], S^t_1 \in (I_0; I_1], \dots, S^t_j \in (I_{j-1}; 1]$:

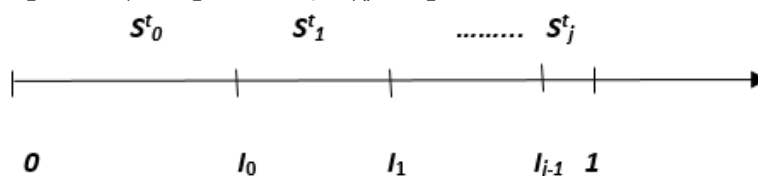


Рис. 2. Области различия состояний S^t_j ресурса R_j

Без потери общности далее будем рассматривать два возможных состояния ресурса – S_0, S_1 . Предположим, что область с номером «0» обозначает нормальное состояние ресурса, а область с номером «1» – критическое состояние.

Рассмотрим процесс функционирования автомата на примере бинарного дерева (рис.3):

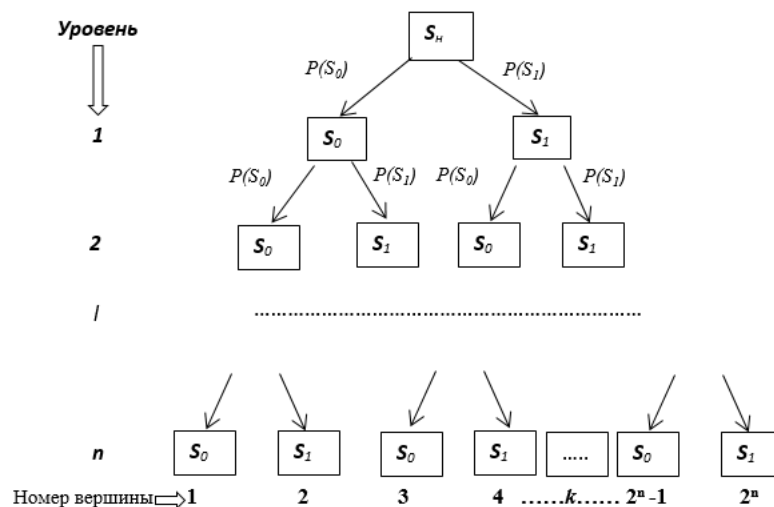


Рис. 3. Структура бинарного дерева функционирования вероятностного автомата.

Пусть $l=1,2,\dots,n$ обозначает уровень дерева, а $k=1,2,\dots,2^n$ – номер вершины на уровне дерева. В качестве входных сигналов будем использовать состояния ресурса в текущий момент времени. На дугах будем отображать вероятности переходов между внутренними состояниями под воздействием входного сигнала. Обозначим $P(S_0)$, $P(S_1)$ – оценки вероятностей переходов в состояния S_0 , S_1 соответственно, $P(S_0) + P(S_1) = 1$. Первоначально автомат находится в начальном состоянии – S_n .

В каждый такт t_i ($i=1,2,\dots,n$) автомат переходит в новое состояние, которое на дереве располагается в соседнем уровне. Пусть текущая координата вершины на дереве обозначена $(l, k(l))$. Тогда автомат под воздействием входного сигнала с уровня l перейдет на уровень $(l+1)$ либо в левую ветвь – состояние S_0 , либо в правую – состояние S_1 . Соответственно новое положение на уровне $(l+1)$ будет определяться как: номер_левый_ $k(l+1) = 2k(l) - 1$, номер_правый_ $k(l+1) = 2k(l)$. Таким образом имеется возможность сохранять последовательность переходов состояний автомата. Эта последовательность представляет собой путь за время n тактов, где n – длина пути на дереве. Путь может быть представлен двоичным словом $d(n)$ длиной n бит, в котором каждый бит $d_i \in \{0,1\}$, где «0» обозначает состояние S_0 , «1» – состояние S_1 . При четной длине пути в двоичном слове встречаются комбинации с одинаковым числом нулевых и единичных бит. Это в ряде случаев может создать дополнительные неопределенности при принятии решений в процессе классификации информационных состояний ресурсов. Поэтому рекомендуется использовать нечетное n .

Таким образом, состояние двоичного слова описывает единственный путь на дереве за n тактов. При равновероятном выборе состояний на уровне l , вероятность выбора любого состояния (вершины) равна $1/2^l$, а вероятность появления пути длиной n определится как $\prod_{i=1}^n (1/2^i)$. Поскольку в общем случае переходы не равновероятны, то оценка вероятности $P(S_{n,k(n)})$ перехода в k -ую вершину на уровне n (обозначим $k(n)$) будет равна

$$P(S_{n,k(n)}) = \prod_{i=1}^n P(S_{i,k(i)})$$

В рабочем режиме функционирования БТС подвержен влиянию внешних воздействий, которые приводят к изменению значений априорных вероятностей. С целью компенсации влияния внешних факторов предлагается на основе апостериорной информации, получаемой в процессе контроля состояния ресурсов БТС, использовать метод, базирующийся на вероятностном адаптивном классификаторе. Главная задача состоит в получении оценок значений условных распределений, определяющих вероятность принадлежности наблюдения каждому из возможных классов. Один из

известных способов основан на применении теоремы Байеса. В контексте решаемой задачи обнаружения изменений состояния ресурсов формула приобретает следующую интерпретацию:

$$P(S_i | N_{il}) = \frac{P(S_i)P(N_{il} | S_i)}{\sum_{i=1}^n P(S_i)P(N_{il} | S_i)} \quad (2)$$

где S_i , $i=\{0,1\}$ – априорная информация о гипотезе оценок вероятности появления состояний;

N_{il} – апостериорная информация появления числа N_{il} состояний S_i на l -ом уровне;

$P(S_i)$ – оценка вероятности появления состояния S_i ;

$P(S_i | N_{il})$ – оценка условной вероятности наблюдения состояния S_i при появлении N_{il} .

Алгоритм определения значений оценок условных распределений появления состояний, принадлежащих каждому из возможных классов, содержит следующую последовательность действий:

1. Задается начальное распределение вероятностей – P_{nj} , значения n , m .
2. Задаются вероятности переходов P_{ijk} .
3. Разыгрывается вектор состояний S_i и строится последовательность $S_{l,k}(0)$, где $l=1,2 \dots n$, $k=1,2, \dots, 2^n$, строится один путь длиной n .
4. Пункт 3 повторяется m раз, строится m путей.
5. Подсчитывается количество N_{in} появления состояний S_i на уровне n .
6. Находятся оценки вероятностей появления состояний $P_i = N_{in} / m$.
7. В соответствии с формулой (2) определяются оценки условных вероятностей принадлежности наблюдений каждому из возможных классов.

Эксперт (лицо принимающее решение) определяет сценарий проведения экспериментов в диалоговом режиме. План экспериментов включает: задание априорных вероятностей P_{nj} , P_{ijk} , длины пути – n , объема выборки – m . Основная цель – провести моделирование процесса обнаружения изменения состояния ресурсов БТС; на основе апостериорной информации определить оценки вероятностей появления состояний S_0 , S_1 , принадлежащих двум противоположным классам – нормальному и критическому на различных уровнях бинарного дерева; обеспечить поддержку принятия решения при оценке влияния стохастической среды и внесении внешнего воздействия; с помощью критериев непараметрической статистики получить оценки влияния значений n и m на достоверность полученных значений $P(S_i | N_{il})$ и др. Полученные статистические результаты экспериментов используются для выбора значений параметров модели с учетом конкретных условий различия состояний ресурсов.

4. Результаты исследования

В статье рассмотрен алгоритмический подход обнаружения уязвимостей интерфейсов БТС на основе вероятностных автоматов. Предлагаемый метод

ориентирован на обнаружение изменения состояния контролируемых ресурсов БТС: канал связи, процессор, память. Метод базируется на применении адаптивной модели вероятностного автомата и байесовского классификатора. Апостериорная информация о состоянии ресурсов в процессе функционирования БТС используется для компенсации воздействий внешней стохастической среды. Предложенный адаптивный подход приведет к повышению достоверности и оперативности процессов поддержки принятия решений при решении задач обеспечения безопасности объектов критической информационной инфраструктуры «Умный город».

Работа выполнена при частичной поддержке Российского фонда фундаментальных исследований (гранты № 19-29-06015/19 и 19-29-06023/19) и Севастопольского государственного университета в рамках внутреннего гранта №28/06-31.

Литература

1. Ширяев А.Н. Вероятностно-статистические методы в теории принятия решений. 2-изд., новое. М.: МЦНМО, 2014. – 144 с.
2. Информационные технологии для критических инфраструктур: монография / под ред. А.В. Скаткова. Севастополь: СевНТУ, 2012. 306 с.
3. Skatkov A. V., Bryukhovetskiy A. A., Moiseev D. V. Intelligent monitoring system for solving large-scale scientific problems in cloud computing environments/ Information and control systems. №2. 2017. pp.19-25
4. Skatkov A., Bryukhovetskiy A., Moiseev D. Detecting changes simulation of the technological objects' information states // MATEC Web of Conferences. Vol. 224, 2018. 02072 (ICMTMTE 2018). <https://doi.org/10.1051/mateconf/201822402072>
5. Поспелов Д.А. Вероятностные автоматы. М.: Энергия, 1970. – 88 с.
6. Зегжда П.Д. Систематизация киберфизических систем и оценка их безопасности / П.Д. Зегжда, М.А. Полтавцева, Д.С. Лаврова // Проблемы информационной безопасности. Компьютерные системы. 2017. № 2. С. 127–138.
7. Al-kahtani M. S. Survey on security attacks in vehicular ad hoc networks (vanets)// 2012 6th International Conference on Signal Processing and Communication Systems, Dec 2012 – pp. 1–9.
8. Pan L., Zheng X., Chen H. Cyber security attacks to modern vehicular systems// Journal of Information Security and Applications. vol. 36. 2017. – pp. 90–100.
9. Markovitz M., Wool A. Field classification, modeling and anomaly detection in unknown can bus networks// Vehicular Communications. vol. 9. 2017. – pp. 43–52.
10. Nilsson D. K., Larson U. E., Picasso F.A. first simulation of attacks in the automotive network communications protocol flexray// Proceedings of the

International Workshop on Computational Intelligence in Security for Information Systems CISIS'08.Springer. 2009 – pp. 84–91.

11. Top 20 and 200 most scanned ports in the cybersecurity industry// SecurityTrails blog, may 07 2019 securitytrails team, <https://securitytrails.com/blog/top-scanned-ports>
12. Шалыто А. А. Парадигма автоматного программирования //Научно-технический вестник СПбГУ ИТМО. Автоматное программирование. 2008. Вып. 53, с. 3–23.

References

1. Shiryaev A.N. Probabilistic and statistical methods in decision theory. 2-ed., New. M.: MCCNMO, 2014 .– 144 p.
2. Information technology for critical infrastructures: monograph / ed. A.V. Skatkova. Sevastopol: SevNTU, 2012. .– 306 p.
3. Skatkov A.V., Bryukhovetskiy A.A., Moiseev D. V. Intelligent monitoring system for solving large-scale scientific problems in cloud computing environments Information and control systems №2 2017 pp. 19-25
4. Skatkov A., Bryukhovetskiy A., Moiseev D. Detecting changes simulation of the technological objects' information states // MATEC Web of Conferences. Vol. 224, 2018. 02072 (ICMTMTE 2018). <https://doi.org/10.1051/mateconf/201822402072>
5. Pospelov D.A. Probabilistic automata. M.: Energy, 1970. – 88 p.
6. Zegzhda P.D. Systematization of cyberphysical systems and assessment of their safety / P.D. Zegzhda, M.A. Poltavtseva D.S. Lavrova// Problemy informacionnoj bezopasnosti. Komp'yuternye sistemy. 2017. № 2. С. 127–138.
7. Al-kahtani M. S. Survey on security attacks in vehicular ad hoc networks (vanets)// 2012 6th International Conference on Signal Processing and Communication Systems, Dec 2012 – pp. 1–9.
8. Pan L., Zheng X., Chen H. Cyber security attacks to modern vehicular systems// Journal of Information Security and Applications. vol. 36. 2017 – pp. 90–100.
9. Markovitz M., Wool A. Field classification, modeling and anomaly detection in unknown can bus networks// Vehicular Communications. vol. 9. 2017. – pp. 43–52.
10. Nilsson D. K., Larson U. E., Picasso F. A first simulation of attacks in the automotive network communications protocol flexray// Proceedings of the International Workshop on Computational Intelligence in Security for Information Systems CISIS'08.Springer. 2009 – pp. 84–91.
11. Top 20 and 200 most scanned ports in the cybersecurity industry// SecurityTrails blog, may 07 2019 securitytrails team, <https://securitytrails.com/blog/top-scanned-ports>
12. Shalyto A. A. The paradigm of automatic programming //Nauchno-tekhnicheskij vestnik SPbGU ITMO. Avtomatnoe programmirovaniye. 2008. № 53, pp. 3–23.